Docket Number: UPN-4105 (M2351)
Title: Compositions and Methods of Using Capsid
Protein from Flaviviruses and Pestiviruses
Inventors: David B. Weiner and Joo-Sung Yang
Sheet 10 of 30

Sequence Range: 1 to 5864

```
                    >BglII
              10  |   20        30        40        50        60        70        80        90       100
GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT
CTGCCTAGCCCTCTAGAGGGCTAGGGGATACCAGCTGAGAGTCATGTTAGACGAGACTACGGCGTATCAATTCGGTCATAGACGAGGGACGAACACACAA

                                                                      >MfeI
             110       120       130       140       150       160 |     170       180       190       200
GGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCG
CCTCCAGCGACTCATCACGCGCTCGTTTTAAATTCGATGTTGTTCCGTTCCGAACTGGCTGTTAACGTACTTCTTAGACGAATCCCAATCCGCAAAACGC

                                                     >SpeI
    >CMV_promoter                                      |
      |
     210       220       230       240       250       260       270       280       290       300
CTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
GACGAAGCGCTACATGCCCGGTCTATATGCGCAACTGTAACTAATAACTGATCAATAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGTATAT

     310       320       330       340       350       360       370       380       390       400
TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
ACCTCAAGGCGCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGCGGGTAACTGCAGTTATTACTGCATACAAGGGTATCA

                                                                            >NdeI
     410       420       430       440       450       460       470       480 |     490       500
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
TTGCGGTTATCCCTGAAAGGTAACTGCAGTTACCCACCTGATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCATGCGGG

                                                                            >SnaBI
     510       520       530       540       550       560       570       580 |     590       600
CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCA
GGATAACTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATACCCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGT

     610       620       630       640       650       660       670       680       690       700
TCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAA
AGCGATAATGGTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGCCAAACTGAGTGCCCCTAAAGGTTCAGAGGTGGGGTAACTGCAGTT

     710       720       730       740       750       760       770       780       790       800
TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
ACCCTCAAACAAAACCGTGGTTTTAGTTGCCCTGAAAGGTTTTACAGCATTGTTGAGGCGGGGTAACTGCGTTTACCCGCCATCCGCACATGCCACCCTC

               >SacI
                |
          >Ecl136II
            |   |
            |   |                                  >T7_promoter
            |   |                                   |
     810  |  820       830       840       850       860 |     870       880       890       900
GTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGT
CAGATATATTCGTCTCGAGAGACCGATTGATCTCTTGGGTGACGAATGACCGAATAGCTTTAATTATGCTGAGTGATATCCCTCTGGGTTCGACCGATCA

>HindIII                  >PflMI
  |                         |
  |  >Kozak_sequence        |
  |   |                     |
>HindIII_linker_[Split]     |
  ||  |                     |
  ||  |    910       920 |     930       940       950       960       970       980       990      1000
TAAGCTTGCCGCCACCATGGATTGGACTTGGATCTTATTTTTAGTTGCTGCTGCTACTAGAGTTCATTCTTCTAAAAAACCAGGTGGCCCCGGCAAGAGC
ATTCGAACGGCGGTGGTACCTAACCTGAACCTAGAATAAAAATCAACGACGACGATGATCTCAAGTAAGAAGATTTTTTGGTCCACCGGGGCCGTTCTCG
          M  D  W  T  W  I  L  F  L  V  A  A  A  T  R  V  H  S>
          _____SIGEY_____>
                                                                 S  K  K  P  G  G  P  G  K  S>
                                                                 _____WNVCHU*_____>
    _____910_____920_____930__903 TO 1335 OF A.PCV/HB.SIGEY-WNVCHU_0_____980_____990_____1000>
    ____10_____20_____30___4_5 TO 437 OF SIGEY-WNVCHU*____70_____80_____90_____100_>
                   _____10_____1 TO 54 OF SIGEY____40_____50____>

                                                                >ApaI
                                                                 |
                                                           >Bsp120I
                                                             |   |
    1010      1020      1030      1040      1050      1060      1070      1080      1090 |    1100
CGCGCCGTGAACATGCTGAAGCGCGGCATGCCCCGCGTGCTGAGCCTGATTGGCCTGAAGCGCGCCATGCTGAGCCTGATCGACGGCAAGGGCCCCATAC
GCGCGGCACTTGTACGACTTCGCGCCGTACGGGGCGCACGACTCGGACTAACCGGACTTCGCGCGGTACGACTCGGACTAGCTGCCGTTCCCGGGGTATG
  R  A  V  N  M  L  K  R  G  M  P  R  V  L  S  L  I  G  L  K  R  A  M  L  S  L  I  D  G  K  G  P  I>
                                                         WNVCHU*>
    _____1010_____1020_____1030_903 TO 1335 OF A.PCV/HB.SIGEY-WNVCHU_70_____1080_____1090_____1100>
    _____110_____120_____130_____1_5 TO 437 OF SIGEY-WNVCHU*___170_____180_____190_____200_>

          >SfiI
           |
    1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
GCTTCGTGCTGGCCCCTGCTGGCCTTCTTCCGCTTCACCGCCATTGCCCCCACCCGCGCCGTGCTGGACCGCTGGCGCGGCGTGAACAAGCAGACCGCCAT
CGAAGCACGACCGGGACGACCGGAAGAAGGCGAAGTGGCGGTAACGGGGGTGGGCGCGGCACGACCTGGCGACCGCGCCGCACTTGTTCGTCTGGCGGTA
  R  F  V  L  A  L  L  A  P  F  F  R  F  T  A  I  A  P  T  R  A  V  L  D  R  W  R  G  V  N  K  Q  T  A  M>
```

FIG 7

Docket Number: UPN-4105 (M2351)
Title: Compositions and Methods of Using Capsid
Protein from Flaviviruses and Pestiviruses
Inventors: David B. Weiner and Joo-Sung Yang
Sheet 11 of 30

```
        4510      4520      4530      4540      4550      4560      4570      4580      4590      4600
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAG
CACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTCCTGTCATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTC

        4610      4620      4630      4640      4650      4660      4670      4680      4690      4700
AGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAA
TCAACCATCGAGAACTAGGCCGTTTGTTTGGTGGCGACCATCGCCACCAAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTTCCTAGAGTTCTT

        4710      4720      4730      4740      4750      4760      4770      4780      4790      4800
GATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA
CTAGGAAACTAGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTCTAATAGTTTTTCCTAGAAGTGGATCT

        4810      4820      4830      4840      4850      4860      4870      4880      4890      4900
TCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC
AGGAAAATTTAATTTTTACTTCAAAAATTTAGTTAGATTTCATATATACTCATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGTGGATAGAG

                                                >AhdI
                                                  |
        4910      4920      4930      4940      4950      4960      4970      4980      4990      5000
AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATG
TCGCTAGACAGATAAAGCAAGTAGGTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCTCCCGAATGGTAGACCGGGGTCACGACGTTAC

        5010      5020      5030      5040      5050      5060      5070      5080      5090      5100
ATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCT
TATGGCGCTCTGGGTGCGAGTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTCGCGTCTTCACCAGGACGTTGAAATAGGCGGA

        5110      5120      5130      5140      5150      5160      5170      5180      5190      5200
CCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
GGTAGGTCAGATAATTAACAACGGCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGTTGCAACAACGGTAACGATGTCCGTAGCACCACAG

        5210      5220      5230      5240      5250      5260      5270      5280      5290      5300
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTC
TGCGAGCAGCAAACCATACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGGGTACAACACGTTTTTTCGCCAATCGAGGAAG

        >PvuI
          |
        5310      5320      5330      5340      5350      5360      5370      5380      5390      5400
GGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGAT
CCAGGAGGCTAGCAACAGTCTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTATTAAGAGAATGACAGTACGGTAGGCATTCTA

            >ScaI
              |
        5410      5420      5430      5440      5450      5460      5470      5480      5490      5500
GCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC
CGAAAAGACACTGACCACTCATGAGTTGGTTCAGTAAGACTCTTATCACATACGCCGCTGGCTCAACGAGAACGGGCCGCAGTTATGCCCTATTATGGCG

        5510      5520      5530      5540      5550      5560      5570      5580      5590      5600
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA
CGGTGTATCGTCTTGAAATTTTCACGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTAGAATGGCGACAACTCTAGGTCAAGCTACATT

        5610      5620      5630      5640      5650      5660      5670      5680      5690      5700
CCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAA
GGGTGAGCACGTGGGTTGACTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGTCCTTCCGTTTTACGGCGTTTTTTCCCTTATT

                                        >SapI
                                          |
        5710      5720      5730      5740      5750      5760      5770      5780      5790      5800
GGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
CCCGCTGTGCCTTTACAACTTATGAGTATGAGAAGGAAAAAGTTATAATAACTTCGTAAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTAC

        5810      5820      5830      5840      5850      5860
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC
ATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGTGTAAAGGGGCTTTTCACGGTGGACTGCAG
```

Docket Number: UPN-4105 (M2351)
Title: Compositions and Methods of Using Capsid
Protein from Flaviviruses and Pestiviruses
Inventors: David B. Weiner and Joo-Sung Yang
Sheet 12 of 30

```
_____WNVCHU*_____>
_____1110_____1120_____1130_903 TO 1335 OF A.PCV/HB.SIGEY-WNVCHU_70_____1180_____1190_____1200>
_____210_____220_____230_____2_5 TO 437 OF SIGEY-WNVCHU*___270_____280_____290_____300__>

       1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
GAAGCACCTGCTGAGCTTCAAGAAGGAGCTGGGCACCCTGACCAGCGCCATCAACCGCCGCAGCAGCAAGCAGAAGAAGCGCGGCGGCAAGACCGGCATT
CTTCGTGGACGACTCGAAGTTCTTCCTCGACCCGTGGGACTGGTCGCGGTAGTTGGCGGCGTCGTCGTTCGTCTTCTTCGCGCCGCCGTTCTGGCCGTAA
  K  H  L  L  S  F  K  K  E  L  G  T  L  T  S  A  I  N  R  R  S  S  K  Q  K  K  R  G  G  K  T  G  I>
_____WNVCHU*_____>
_____1210_____1220_____1230_903 TO 1335 OF A.PCV/HB.SIGEY-WNVCHU_70_____1280_____1290_____1300>
_____310_____320_____330_____3_5 TO 437 OF SIGEY-WNVCHU*___370_____380_____390_____400__>


                                             >BstEII
                                    >PaeR7I |
                                    |       |
                          >NotI   >XhoI |        >BstBI
                          |         |    |        |
                   >NotI_introduction_[Split]     >V5_epitope
                          ||      ||         |
       1310      1320     1330  ||  1340|   |1350|  |1360   1370      1380      1390      1400
GCCGTGATGATTGGCCTGATCGCCAGCGTGGGCGCGGCCGCTCGAGGTCACCCATTCGAAGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTA
CGGCACTACTAACCGGACTAGCGGTCGCACCCGCGCCGGCGAGCTCCAGTGGGTAAGCTTCCATTCGGATAGGGATTGGGAGAGGAGCCAGAGCTAAGAT
  A  V  M  I  G  L  I  A  S  V  G  A>
_____WNVCHU*_____>
_____903 TO 1335 OF A.PCV/HB.SIGE____>
_____5 TO 437 OF SIGEY-WNVCHU*_____>

   >AgeI                          >PmeI
   |                              |
   >Polyhistidine_tag             |       >BGH_polyadenylation_signal
   |                              |       |
   |1410 |    1420      1430      |1440   1450      1460      1470      1480      1490      1500
CGCGTACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCC
GCGCATGGCCAGTAGTAGTGGTAGTGGTAACTCAAATTTGGGCGACTAGTCGGAGCTGACACGGAAGATCAACGGTCGGTAGACAACAAACGGGGAGGGG

       1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
CGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG
GCACGGAAGGAACTGGGACCTTCCACGGTGAGGGTGACAGGAAAGGATTATTTTACTCCTTTAACGTAGCGTAACAGACTCATCCACAGTAAGATAAGAC

                                             >BbsI
                                             |
       1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA
CCCCCACCCCACCCCGTCCTGTCGTTCCCCCTCCTAACCCTTCTGTTATCGTCCGTACGACCCCTACGCCACCCGAGATACCGAAGACTCCGCCTTTCTT

       1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
CCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAG
GGTCGACCCCGAGATCCCCCATAGGGGTGCGCGGGACATCGCCGCGTAATTCGCGCCGCCCACACCACCAATGCGCGTCGCACTGGCGATGTGAACGGTC

       1810      1820      1830      1840      1850      1860      1870      1880      1890      1900
CGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTC
GCGGGATCGCGGGCGAGGAAAGCGAAAGAAGGGAAGGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCAGTTCGAGATTTAGCCCCGTAGGGAAATCCCAAG

                                             >DraIII
                                             |
       1910      1920      1930      1940      1950      1960|   1970      1980      1990      2000
CGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGA
GCTAAATCACGAAATGCCGTGGAGCTGGGGTTTTTTGAACTAATCCCACTACCAAGTGCATCACCCGGTAGCGGGACTATCTGCCAAAAAGCGGGAAACT

       2010      2020      2030      2040      2050      2060      2070      2080      2090      2100
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGGG
GCAACCTCAGGTGCAAGAAATTATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGAGCCAGATAAGAAAACTAAATATTCCCTAAAACCC

       2110      2120      2130      2140      2150      2160      2170      2180      2190      2200
GATTTCGGCCTATTGGTTAAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGAAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAG
CTAAAGCCGGATAACCAATTTTTTACTCGACTAAATTGTTTTTAAATTGCGCTTAATTAAGACACCTTACACACAGTCAATCCCACACCTTTCAGGGGTC

       2210      2220      2230      2240      2250      2260      2270      2280      2290      2300
GCTCCCCAGGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAA
CGAGGGGTCCGTCCGTCTTCATACGTTTCGTACGTAGAGTTAATCAGTCGTTGGTCCACACCTTTCAGGGGTCCGAGGGGTCGTCCGTCTTCATACGTTT

       2310      2320      2330      2340      2350      2360      2370      2380      2390      2400
GCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTG
CGTACGTAGAGTTAATCAGTCGTTGGTATCAGGGCGGGGATTGAGGCGGGTAGGGCGGGGATTGAGGCGGGTCAAGGCGGGTAAGAGGCGGGGTACCGAC

                                                                                   >AvrII
                                                                                   |
                                                                                   >StuI
                                                                                   ||
       2410      2420      2430      2440      2450      2460      2470      2480  ||  2490      2500
ACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAA
TGATTAAAAAAAATAAATACGTCTCCGGCTCCGGCGGAGACGGAGACTCGATAAGGTCTTCATCACTCCTCCGAAAAAACCTCCGGATCCGAAAACGTTT

   >SmaI
   |
   >XmaI                                      >BsaBI
   |                                          |
   |2510      2520      2530      2540      2550  | 2560      2570      2580      2590      2600
AAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTC
TTCGAGGGCCCTCGAACATATAGGTAAAAGCCTAGACTAGTTCTTCTGTCCTACTCCTAGCAAAGCGTACTAACTTGTTCTACCTAACGTGCGTCCAAGAG
```